DECIPHERING BIOSYNTHETIC MACHINERY OF PODOPHYLLOTOXIN- ANTICANCER METABOLITE OF *PODOPHYLLUM HEXANDRUM* ROYLE

SYNOPSIS

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

In

Biotechnology

By

PAWAN KUMAR

Enrolment No. 126569

Under the supervision of

Prof. (DR.) Rajinder Singh Chauhan (Supervisor)

DR. Hemant Sood (Co-supervisor)

DEPARTMENT OF BIOTECHNOLOGY AND BIOINFORMATICS

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY

WAKNAGHAT-173234, HP, INDIA

FEBRUARY 2016

DECIPHERING BIOSYNTHETIC MACHINERY OF PODOPHYLLOTOXIN- ANTICANCER METABOLITE OF *PODOPHYLLUM HEXANDRUM* ROYLE

PAWAN KUMAR

Introduction

Podophyllum hexandrum Royle (*P. hexandrum*) is native to alpine and sub-alpine regions of the North-Western Himalayas, North America and South West of China (2500-4200m) and commonly used as a source of podophyllotoxin and its derivatives- the anticancer bioactive compounds. Podophyllotoxin is present in different plant species *Podophyllum hexandrum* Royle (*syn. P. emodi* Wall), *Podophyllum peltatum, Linum album, Callitris drummondii, Juniperous chinensis, Polygala polygama*, etc., however *P. hexandrum* has been used as a primary source due to higher content (4.3%) ([1,2]. Several anticancer drugs such as Etoposide (VP-16-213), VUMON® (Teniposide; VM-26), Etopophos®, Pod-Ben-25, Condofil, Verrusol and Warticon prepared from podophyllotoxin and its derivatives are available in the market. Many other drugs, GL331, Top 53, NK 611 and CPH 82 are in phase III clinical trials containing podophyllotoxin and its derivatives [3]. Owing to these medicinal properties, *P. hexandrum* has been overexploited and therefore, placed under the category of critically endangered plant species [4].

Podophyllotoxin is commercially isolated from roots and rhizomes of *Podophyllum* species. However, limited availability and finite amount of raw material necessitate alternate sources/production routes for meeting increasing demands of podophyllotoxin. Different cell cultures interventions have been tried for the production of podophyllotoxin through cell cultures systems such as elicitor's treatments, precursor feeding, modification in production media, immobilization of cells and scale-up through bioreactors, etc. [5]. However, podophyllotoxin yields have been low (0.3%) in callus/cell suspension cultures as compared to roots and rhizomes of plants grown in natural habitat [6, 7]. Further increase in podophyllotoxin content can only be achieved through genetic interventions in podophyllotoxin biosynthetic machinery. However, knowledge of podophyllotoxin biosynthetic pathway is fragmentary with missing links in the pathway as well as lack of relative contribution of pathway steps in contributing to the podophyllotoxin content. Podophyllotoxin and its derivatives are derived from the shikimic acid/phenylpropanoid

pathway which has not yet been completely elucidated [1, 8]. Till date, only 31 steps are known in the podophyllotoxin biosynthetic pathway up to the intermediate compound pluviatolide; thereafter possible 36 steps were predicted but not validated either for the structures of intermediates or for genes/proteins/enzymes [9]. Total 26 genes encoding podophyllotoxin pathway enzymes are known to catalyse 31 known steps, starting from D-erythrose-4-phosphate to intermediate compound pluviatolide (Fig.1). Out of 26 genes, only 7 genes, PS, PLR, SD, DPO, CAD, C4H and PCH have been characterized in *Podophyllum* spp. to correlate their role in podophyllotoxin biosynthesis through expression analysis [10].

No information exists as of today on regulatory factors (key pathway genes, transcription factors (TFs), ABC transporters and miRNAs) controlling the biosynthesis and accumulation of podophyllotoxin content in *P. hexandrum*. Therefore, studies must be undertaken to elucidate or decipher important molecular components controlling podophyllotoxin content in *P. hexandrum* so that metabolic engineering can be undertaken to fulfil escalating demand of podophyllotoxin.

The availability of genomics, transcriptomics and metabolomics resources can provide a powerful tool for identification, validation, characterization and functional analyses of molecular components in order to infer their role in regulation of metabolic processes in plant species. The assembled transcriptomes data of *Podophyllum* species (*P. hexandrum*—2, 27, 885 contigs, *P. peltatum*—1, 47, 960 contigs) were downloaded from the medicinal plants transcriptomics database (http://medplants.ncgr.org/) [9] and the transcriptomes were mined and analyzed for identification, characterisation and classification of molecular components by using in-house developed PERL scripts. This resource substantially broadens our understanding of podophyllotoxin biosynthesis in *P. hexandrum*.

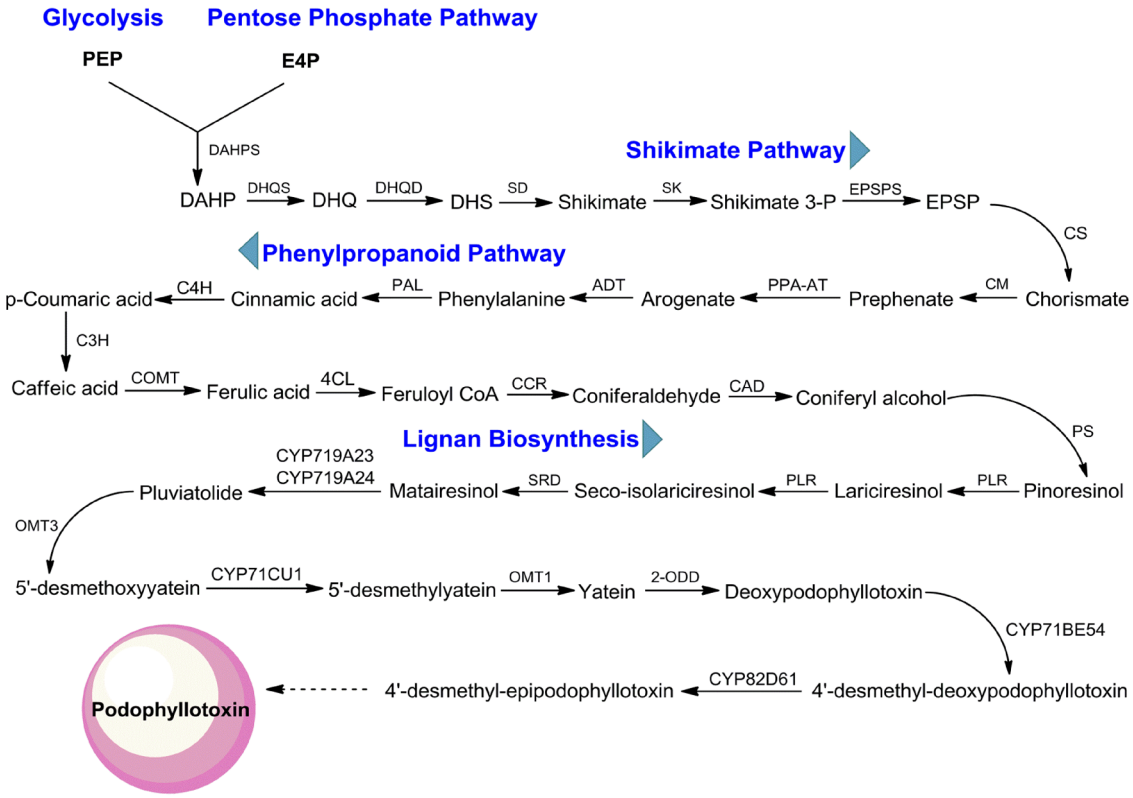

Fig. 1 Schematic representation of podophyllotoxin biosynthetic pathway [1, 8, 9]

The current research work focussed on identification, characterisation, validation and expression analysis of molecular components such as biosynthetic pathway genes, transcription factors, ABC transporters and miRNAs to deduce the biosynthetic machinery of podophyllotoxin in *P. hexandrum.* The current work revealed crucial candidate pathway genes, key TFs, ABC transporters and miRNAs associated with podophyllotoxin biosynthesis through extensive transcriptome analysis followed by wet lab experiments.

Keeping in view the immense medicinal importance of *P. hexandrum*, and lack of knowledge on molecular basis of podophyllotoxin biosynthetic machinery, the present work was carried with the following objectives:

Objective 1: Identification of pathway components/genes contributing to podophyllotoxin biosynthesis in *Podophyllum hexandrum*

Objective 2: Dissecting regulatory machinery/components of podophyllotoxin biosynthesis and accumulation in *Podophyllum hexandrum*

Materials and Methods

Plant materials and growth conditions

Eleven accessions of *P. hexandrum* Royle belonging to different geographical locations of North-Western Himalayas were procured from the Himalayan Forest Research Institute at Jagatsukh (altitude1900m asl), Kullu, Himachal Pradesh, India. The plant were maintained in plastic pots (20 cm height x 20 cm top width x 12 cm bottom width) containing soil, sand and farm yard manure mixture in a ratio of 2:1:1 with controlled environmental conditions [light (intensity 300-1400 lx), temperature (25±2^{0}C), relative humidity (≈ 80%) and photoperiod (14 h day/10h night) at the green house of Jaypee University of Information Technology, Waknaghat, Solan, India. Shoot, root and rhizome tissues of *P. hexandrum* (3 year old plants) were harvested, frozen in liquid nitrogen and stored at -80^{0}C for metabolite quantification and expression analysis through qRT-PCR.

Quantification of podophyllotoxin content

The quantification of podophyllotoxin content (expressed as % of fresh wt) was carried by reverse phase high-performance liquid chromatography (HPLC) in shoots, roots and rhizomes tissues of eleven accessions of *P. hexandrum*. The quantification of podophyllotoxin was performed on Waters HPLC system (HPLC Waters 515) equipped with 515 HPLC Pump, Waters 717 autosampler, Waters 2996 photodiode-array detector and Waters Spherisorb® C18 (4.6×250 mm, 5 µm) reverse phase column using Empower-2 software at 210 nm UV wavelength [11]. The podophyllotoxin was identified on the basis of retention time and comparison of UV spectra with the authentic podophyllotoxin standard (Chroma Dex, CA, USA).

Analysis of transcriptomes of *Podophyllum* species

Assembled transcriptomes data of *Podophyllum* species (*P. hexandrum* - 2, 27, 885 contigs, *P. peltatum*- 1, 47, 960 contigs) were downloaded from medicinal plants transcriptomics database (http://medplants.ncgr.org/) [9] and the transcriptomes were mined and analyzed for molecular components such as biosynthetic pathway genes, transcription factors (TFs), ABC transporters and miRNAs by using in-house developed PERL scripts. Workflow depicting mining, identification and experimental validation of molecular components from *Podophyllum* species transcriptomes has been given in Fig. 2.

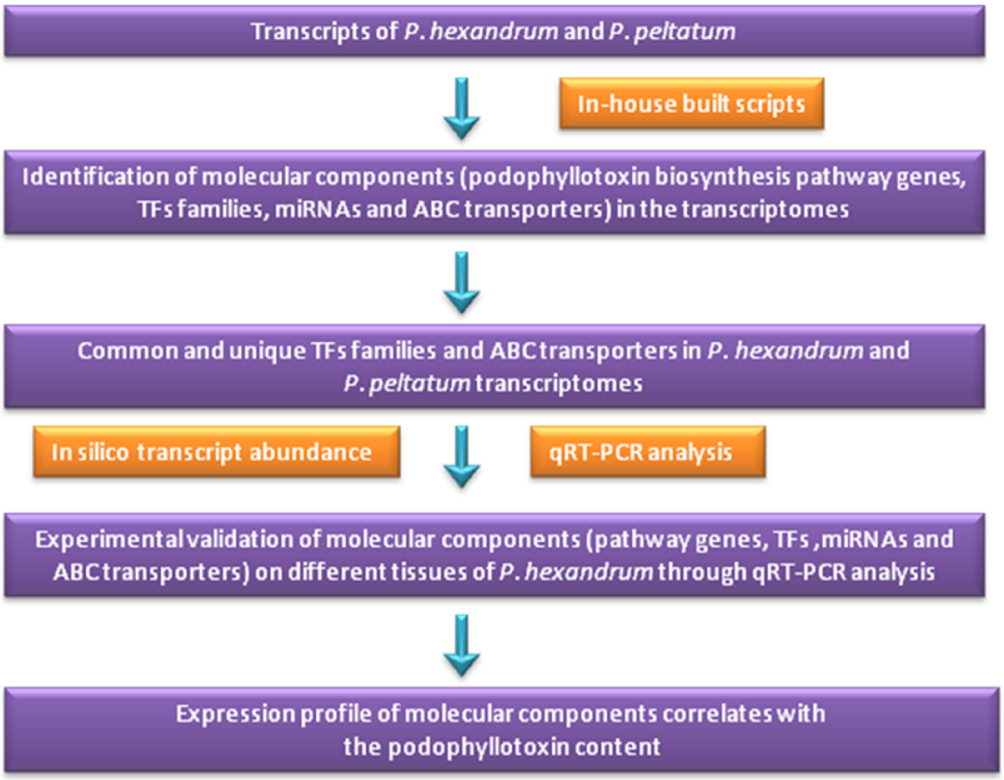
Transcripts of P. hexandrum and P. peltatum
In-house built scripts
Identification of molecular components (podophyllotoxin biosynthesis pathway genes, TFs families, miRNAs and ABC transporters) in the transcriptomes
Common and unique TFs families and ABC transporters in P. hexandrum and P. peltatum transcriptomes
In silico transcript abundance
qRT-PCR analysis
Experimental validation of molecular components (pathway genes, TFs ,miRNAs and ABC transporters) on different tissues of P. hexandrum through qRT-PCR analysis
Expression profile of molecular components correlates with the podophyllotoxin content

Fig. 2 Workflow depicting mining, identification and experimental validation of molecular components from *Podophyllum* species transcriptomes

Identification of pathway components/genes contributing to podophyllotoxin biosynthesis

The functional annotation of *P. hexandrum* transcriptome was done by comparing transcripts against the NR protein sequences database available at NCBI by using BLASTX algorithm with E-value threshold of 10^{-05}. Using in-house developed perl scripts, fourteen genes (*DAHPS, DHQS, DHQD, COMT, EPSPS, CS, CM, PPA-AT, PD, ADT, PAL, 4CL, COMT, CCR*) related to podophyllotoxin biosynthesis were identified based on high homology (identity value ≥ 30%). Besides this, nucleotide sequences of 7 genes, namely *CYP719A23, PLR, SRD, PS, CAD, C4H,* and *C3H* were retrieved from *P. hexandrum* CDSs available at GenBank Database (http://www.ncbi.nlm.nih.gov/genbank/).

Computational mining of TFs families in transcriptomes of *Podophyllum* species

Plant transcription factor database (PlantTFcat) (http://plantgrn.noble.org/PlantTFcat/familylist.do) was used for identification and

classification of TFs families involved in secondary metabolites biosynthesis. Transcripts coding for TFs families in transcriptomes of *Podophyllum* species (*P. hexandrum* and *P. peltatum*) were identified by searching against all transcription factor protein sequences at PlantTFcat using BLASTX with an E-value cutoff $1e^{-6}$ and according to the presence of InterPro domains. All transcripts were annotated to proteins (according to possible open reading frames) and searched through InterProScan. PlantTFcat classified these transcripts into 108 known TF families. The conserved domains in TFs were identified by using conserved domain database available at NCBI (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).

Mining ABC transporters in transcriptomes of *P. hexandrum* and their phyllogenetic analysis

P. hexandrum assembled transcriptomes data (2, 27, 885 contigs) were downloaded from the medicinal plants transcriptomics database (http://medplants.ncgr.org/, Marques *et al.* 2013) and the transcriptomes were mined and analyzed for ABC transporters by using in-house developed PERL scripts. Protein family database (Pfam) (pfam.xfam.org) was used for identification of ABC transporters in the transcriptomes of *P. hexandrum*. Transcripts encoding ABC transporters in *P. hexandrum* were mined by searching the transcriptomes against all available ABC transporters sequences at Pfam database using in-house developed perl scripts. For phylogenetic reconstruction of identified ABC transporters, protein sequences of known ABC transporters characterized in model plant species were downloaded from NCBI Genbank database. The alignment of sequences was carried out by CLUSTALX and final tree was constructed through MEGA 6 (Molecular Evolutionary Genetics Analysis 6.0) software using Maximum Likelihood method [12].

Mining miRNAs and their mRNAs targets in transcriptomes of *P. hexandrum*

miRNAs were identified from transcriptomes of *P. hexandrum* using in-house built perl-scripts. PMRD (plant microRNA database) (http://bioinformatics.cau.edu.cn/PMRD/) was used as a reference database for mining conserved miRNAs in different tissues/organs of *P. hexandrum*. psRNATarget (A Plant Small RNA Target Analysis Server) (http://plantgrn.noble.org/psRNATarget/) target prediction tool was used to identify miRNAs targets in P. *hexandrum* transcriptomes [13]. NCBI-NR database was used for identification of functions governed by miRNAs by BLASTX program with default parameters. The

identified miRNAs targets were involved in various physio-biological processes in *P. hexandrum*. In order to better understand the functions of miRNAs, their possible targets were identified through GO, COG and KAAS using default parameters.

In-silico transcript abundance for pathway genes, ABC transporters and transcription factors through RSEM

RSEM was used for quantification of transcript abundance from the RNA-Seq high quality data. RSEM measures the expression profiles interms of FPKM (Fragments per kilobase of transcript per million mapped reads) and RPKM (Reads per kilobase per million) [14]. The transcript abundance of the transcripts encoding pathway genes, ABC transporters and transcription factors from the transcriptomes of *P. hexandrum* was calculated using FPKM parameter of RSEM package. All default parameters were used in the query option.

Identification of putative *cis*-regulatory elements

The *cis*-regulatory elements in the promoter regions of biosynthetic pathway genes were identified through *in-silico* analysis using Plant Genome Database (http://www.plantgdb.org/) and Plant *cis*-acting Regulatory DNA Elements Database (PLACE) (https://sogo.dna.affrc.go.jp/cgiin/sogo.cgi?lang=en&pj=640&action=page&page=newplace) [15]. The complete coding sequences of pathway genes (*CAD, C4H, PS, C3H*, *PAL, 4CL, PLR* and *SRD*) were retrieved from NCBI Genbank and BlastN analysis was performed in Plant Genome Database (http://www.plantgdb.org/) against the genome sequences of *Arabidopsis thaliana*, *Solanum tubersoum*, *Vitis vinifera* and *Glycine max*. The *cis*-regulatory elements were mined through PLACE database by taking upstream region up to 1 kb from translation start site.

Isolation of RNA and cDNA synthesis

Total RNA was isolated from 100 mg of frozen tissues of shoots and rhizomes of *P. hexandrum* by using QIAGEN RNeasy mini kit and digested with DNaseI (Invitrogen). First-strand complementary DNA synthesis (cDNA) was done using Verso cDNA synthesis kit (Thermo Scientific, USA) from total RNA (2 μg) template as per manufacturer's instructions. The primers for qRT-PCR analysis were designed using Primer 3 v.0.4.0 software.

Experimental validation and expression analysis using quantitative real-time PCR analysis

Quantitative real-time PCR (qRT-PCR) based analysis for biosynthetic pathway genes, transcription factors, ABC transporters and miRNAs were performed on shoots, roots and rhizome tissues of *P. hexandrum*. PCR reaction was performed in 12.5 µl mixture containing diluted cDNA sample (100 ng) as a template, SYBR Green PCR master mix (Bio-Rad Laboratories) and 10µM each of forward and reverse gene-specific primers (Table 1). The reactions were performed using CFX96 system (Bio-Rad Laboratories; Hercules CA, USA) with the following programme: 94°C (3 min), 94°C (30 s), corresponding annealing temperature (30 s), 72°C (20 s) at 40 cycles. To normalize the variance in cDNA input, *cyclophilin*, *Elongation factor-1 alpha* (*EF-1 alpha*) and 5.8sRNA genes were used as an internal control. The relative expression of each gene was calculated using the comparative Ct value method. The significant differences between transcripts were statistically evaluated by standard deviation.

Results

Podophyllotoxin content

Highest amount of podophyllotoxin was detected in rhizomes (3.05%) followed by roots (1.37%) and shoots (0.01%) of *P. hexandrum* (Fig.3). The podophyllotoxin content differed in different tissues/organs, thereby providing differential conditions to correlate the transcript abundance of molecular components (pathway genes, TFs, ABC transporters and miRNAs) with podophyllotoxin content in *P. hexandrum*. To add one more filter to further confirm the correlation between transcripts profile and metabolite content, we also quantified podophyllotoxin content in rhizomes of 11 different accessions of *P. hexandrum* collected from different geographical locations. Podophyllotoxin content ranged from 0.59% to 3.05% among different accessions (Fig. 4). The accession Ph5 of *P. hexandrum* showed the highest amount of podophyllotoxin (3.05%) whereas accession Ph6 contained lowest podophyllotoxin content (0.59%). The high and low podophyllotoxin content aceesions of *P. hexandrm* were selected for the expression analysis through qRT-PCR. The podophylltoxin conent was not detected in shoots of these accessions.

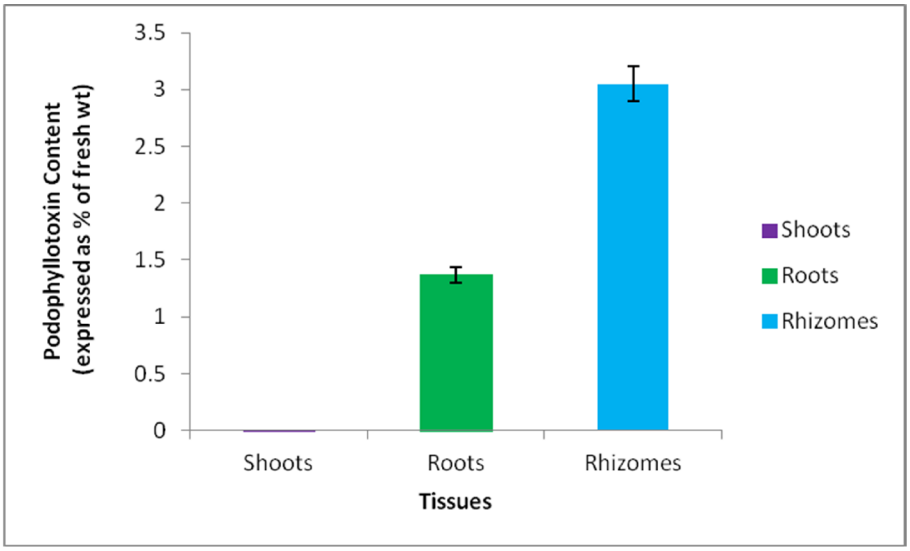

Fig. 3 Podophyllotoxin content in different tissues of *P. hexandrum*

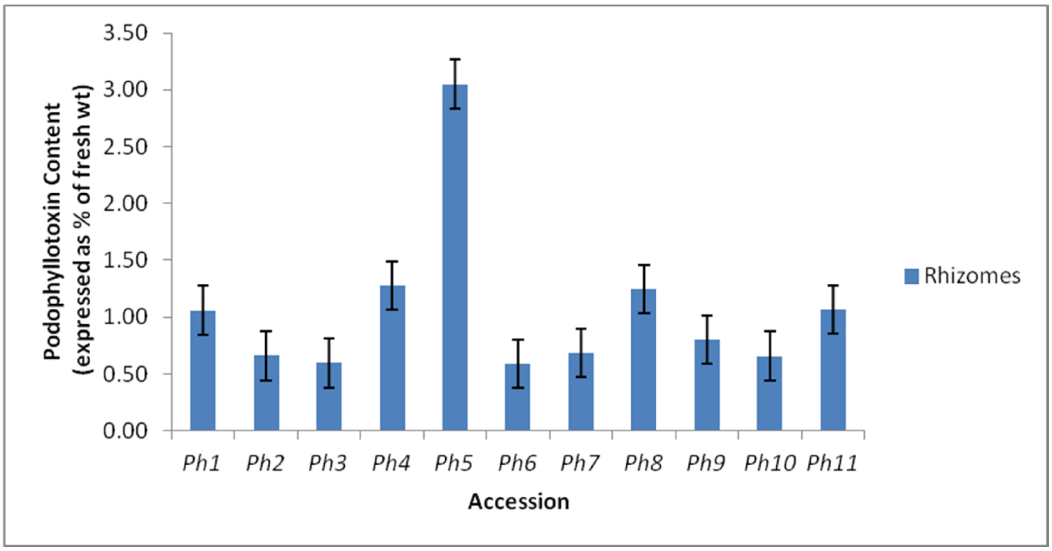

Fig. 4 Podophyllotoxin content in different accessions of *Podophyllum hexandrum*

Identification of podophyllotoxin biosynthetic pathway components/genes through transcriptome mining

The whole picture of podophyllotoxin biosynthetic pathway has not yet been completely elucidated. Till date, 31 steps were known up to intermediate compound pluviatolide in podophyllotoxin biosynthesis; thereafter it is unknown until the end product podophyllotoxin. Fourteen genes namely, *DAHPS, DHQS, DHQD, COMT, EPSPS, CS, CM, PPA-AT, PD, ADT, PAL, 4CL, COMT, CCR* involved in the podophyllotoxin biosynthetic pathway were

identified in the transcriptomes of *P. hexandrum* on the basis of high sequence identity (≥ 30% and E-value threshold of 10^{-05}) with the genes in other plant species. Four genes, *3DSD, QD, SK* and *PA4H* could not be ascertained due to their low sequence similarity (Fig. 1).

Identification of TFs families in transcriptomes of *Podophyllum* species

The mining of *P. hexandrum* and *P. peltatum* transcriptomes derived from shoots and rhizomes resulted in the identification of 3477, 3517 and 3383, 3824 TFs encoding transcripts belonging to 97 and 95 TF families. The identified TFs encoding transcripts were suggested to regulate different biological processes such as photosynthesis, seed germination, disease resistance, secondary metabolites biosynthesis, etc (https://sites.google.com/site/combiogroup/data-access). We further identified TFs encoding transcripts contributing to the biosynthesis of secondary metabolites in *Podophyllum* species. Literature survey showed that eleven TFs families namely, WRKY, MYB, bZIP, bHLH, NAC, DOF, Zing finger, SPL, AP2/ERF, mTERF and TFIIIA have been implicated in biosynthesis of secondary metabolites in various plant species [16, 17]. Out of eleven, four TFs families, WRKY, MYB, bZIP and bHLH have been suggested to contribute in lignin and terpenoid biosynthesis. Therefore, we mined transcriptomes of *Podophyllum* species for these four TFs families which revealed 82, 278, 70 and 90 transcripts in shoots and 89, 273, 72 and 91 transcripts in rhizomes of *P. hexandrum* transcriptomes and similarly, 70, 268, 48 and 92 transcripts in shoots and 58, 245, 41 and 85 transcripts in rhizomes of *P. peltatum* transcriptomes corresponding to bZIP, MYB, WRKY and bHLH TF families (Fig. 5). The identified TFs families are suggested to be involved in the regulation of podophyllotoxin biosynthesis. The multiple sequence alignment of TF encoding transcripts was performed using clustalW and corresponding conserved domains were identified through conserved domain database available at NCBI (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).

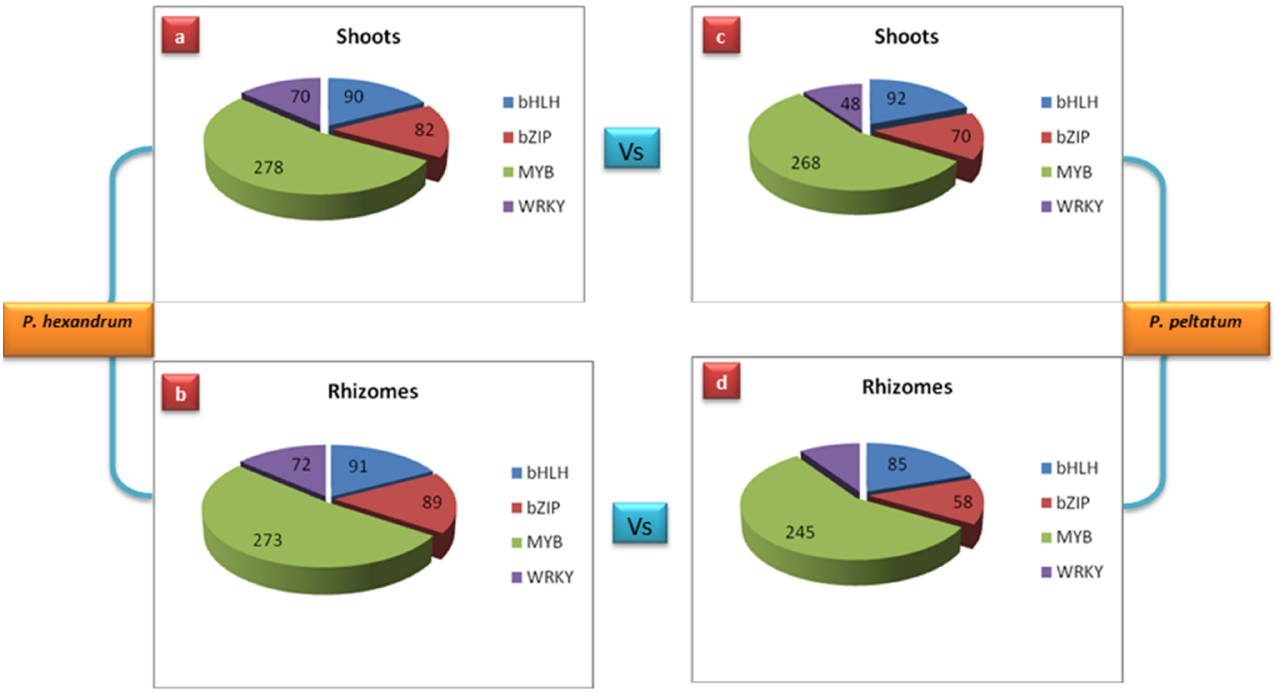

Fig. 5 Distribution of four TFs families in shoots and rhizomes transcriptomes of *Podophyllum* species a) & b) *P. hexandrum;* c) & d) *P. peltatum*

Mining ABC transporters from transcriptomes of *P. hexandrum* and phyllogenetic analysis

Total 252 putative ABC transporters were identified in the transcriptome of *P. hexandrum* by searching protein sequences corresponding to full complement of ABC transporters through Pfam database. The identified transporters genes were suggested to be involved in the transport of primary and secondary metabolites in *P. hexandrum.* We further identified orthologous of five functionally validated transporters genes, three from *Arabidopsis thaliana* (*PhABCI, PhABCII, PhABCIII*), and one each from *Vitis vinifera* (*PhABCIV*) and *Nicotiana plumbaginifolia* (*PhABCV*) in the transcriptomes of *P. hexandrum*. These transporter genes are assumed to play a role in the transport of secondary metabolites in *P. hexandrum* as has been demonstrated in other plant species. Phylogenetic analysis of identified ABC transporters was performed through MEGA6 software using functionally characterised ABC transporters retrieved from model plant species such as *Arabidopsis thaliana*, *Nicotiana plumbaginifolia, Zea mays* and *Medicago truncatula.* The ABC transporters genes showed

high homology with well characterised genes in model plants that further confirms the biological functions of identified ABC transporters in *P. hexandrum*.

Experimental validation and differential expression analysis of TFs and ABC transporters using qRT-PCR in *P. hexandrum*

The transcripts of four TFs were validated through qRT-PCR in *P. hexandrum* tissues differing for podophyllotoxin content. The partial cloned fragments of TFs matched transcriptome contigs sequences perfectly (100%), therefore, again validating the transcriptomes mining procedures employed in the present study. The partial sequences of TFs (212-434bp) have been submitted to the GenBank under accession numbers KP299959 to KP299961. These sequences were utilized for designing primers for gene expression analysis. *In-silico* transcript abundance values of respective TFs ranged between 0.0014-12.01 in shoots and rhizomes of *P. hexandrum*. In order to ascertain the relevance of FPKM-based expression values obtained through RSEM analysis, qRT-PCR was carried out in shoots and rhizomes of *P. hexandrum.* Increase in transcript abundance with 1.09-folds for MYB, 0.62-folds for WRKY and 0.26-folds for bHLH were observed in rhizomes with respect to shoots of *P. hexandrum.* The bZIP class of TFs showed higher transcript abundance with 19.60-folds in rhizomes compared to shoots of *P. hexandrum* (Fig. 7). The 3 TFs belonging to bZIP, MYB and WRKY families showed down-regulation in transcript amounts in shoots compared to rhizomes, except for bHLH TF with higher amount of transcript in shoots compared to rhizomes.

Furthermore, the transcript abundance of identified transcription factors was further assessed between high (Ph5) (3.05%) versus low (Ph6) (0.59%) podophyllotoxin content accessions of *P. hexandrum*. This comparative expression analysis further confirmed the expression status of transcription factors in rhizomes of *P. hexandrum* (Fig. 7). Transcript abundance of TFs obtained through two different platforms i.e. RSEM and qRT-PCR resulted in the identification of potential TFs possibly regulating podophyllotoxin biosynthesis in *P. hexandrum*. The Pearson's correlation coefficient was performed to find out relationship between podophyllotoxin content and transcript abundance of TFs. The transcript abundance pattern of TFs (bZIP, MYB, WRKY and bHLH) showed positive correlation with podophyllotoxin content which suggested their role in regulating and controlling the podophyllotoxin biosynthesis in *P. hexandrum*. The bZIP TFs showed highest correlation ($\geq$ 0.99) with podophyllotoxin biosynthesis.

In case of transporters, seventeen ABC transporters genes were selected on the basis of higher transcript abundance value for validation through qRT-PCR in shoots, roots and rhizome tissues varying for podophyllotoxin contents. The qRT-PCR-based relative transcript abundance of transporters genes ranged from 0.002 to 1.673 in shoots, roots and rhizomes of *P. hexandrum* (Fig. 8). Gene expression pattern observed through two different approaches was in agreement with each other. The *PhABC3*, *PhABC6* and *PhABCIV* transporter genes showed highest transcript abundance with 1.673, 1.25 and 1.0 and in shoots, roots and rhizomes of *P. hexandrum*. Most of the ABC transporters genes showed higher transcript abundance in shoots suggesting their role in transporting macro- and micro- molecules in shoots. Moreover, the expression of five functionally characterized transporters genes (*PhABCI*, *PhABCII*, *PhABCIII*, *PhABCIV* and *PhABCV*) from *Arabidopsis thaliana, Vitis vinifera* and *Nicotiana plumbaginifolia* were also evaluated through qRT-PCR across different tissues/organs of *P. hexandrum.* Expression analysis revealed that *PhABCII*, *PhABCV* showed higher transcript abundance in shoots and *PhABCIII* in rhizomes of *P. hexandrum*.

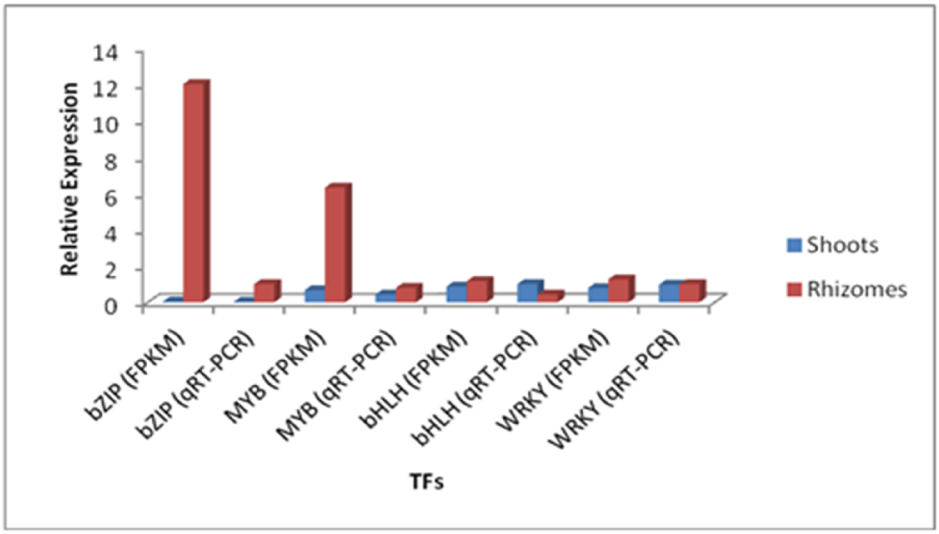

Fig. 6 Relative expression analysis of TF genes in shoots and rhizomes of *P. hexandrum* based upon the data obtained by fragments per kilobase of exon per million fragments mapped (FPKM) values and validated by quantitative real time-polymerase chain reaction (qRT-PCR)

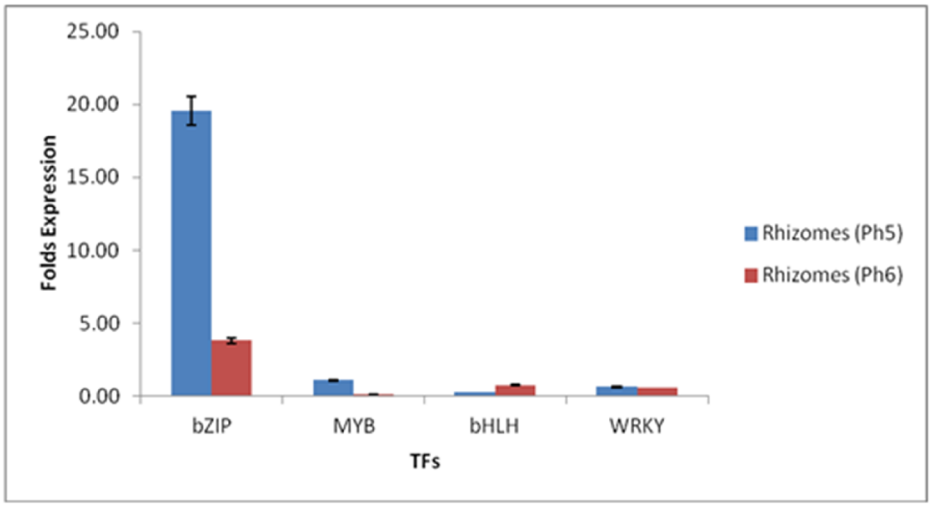

Fig. 7 Expression status of transcription factors in rhizomes of high (Ph5) versus low (Ph6) podophyllotoxin content accessions of *P. hexandrum*

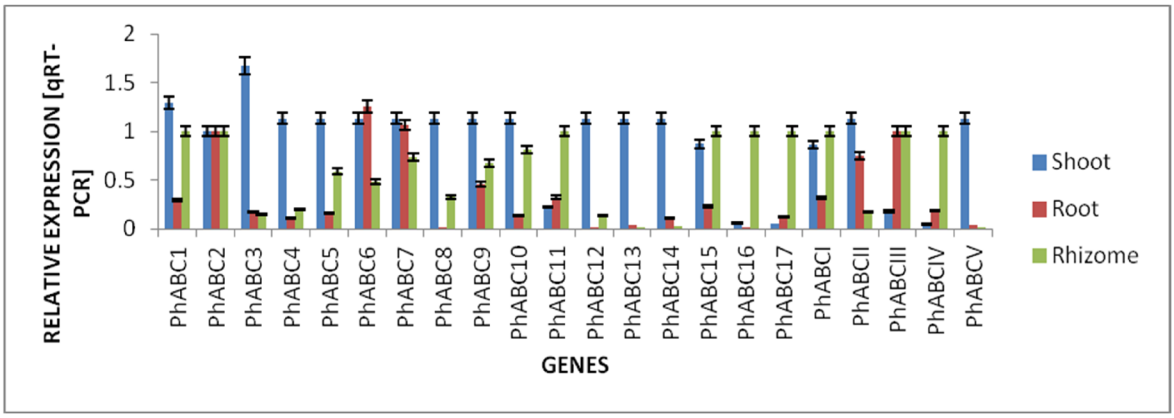

Fig. 8 Relative expression of transporter genes in shoots, roots and rhizomes of *P. hexandrum* analyzed by qRT-PCR. *Error bars* indicate percentage error (5 %)

Identification of miRNAs and their mRNAs targets in transcriptomes

PMRD (plant microRNA database) (http://bioinformatics.cau.edu.cn/PMRD/) was used to identify miRNAs from transcriptomes of *P. hexandrum* [18]. Precursor miRNAs sequences were mined and used for identifying potential miRNAs using in-house perl scripts and the criteria are mentioned in the materials and method section. Total 279 putative miRNAs were identified from transcriptomes that were conserved across various plant species ranging in size from 18 to 22 nucleotides with perfect complementarity (100%*)* to homologous

sequences. Out of 279, 6 miRNAs were selected for further validation on the basis of their possible role in regulation of podophyllotoxin biosynthesis.

Total 375 targets were found for 6 conserved miRNAs in *P. hexandrum* having perfectly or near-perfect complementarity to their mRNA targets using psRNATarget tool. The identified potential targets were involved in various biological processes including, plant defense response such as biotic and abiotic stresses, secondary metabolisms, growth and development, signal transduction, host–pathogen interaction etc. The miRNA miR2673a/b has maximum number of targets (186 targets). The functional analysis of identified putative targets of miRNAs was performed through KAAS (KEGG automatic annotation server) (http://www.genome.jp/tools/kaas/) using default settings [19]. KAAS provides information pertaining to metabolism, genetic information processing and signaling and cellular processes. Total 26 targets were found for secondary metabolites biosynthesis through KAAS. Out of 26, 8 targets were selected for further validation and expression profiling on the basis of their association with podophyllotoxin biosynthesis.

Experimental validation and differential expression profiling of miRNAs using stem-loop qRT-PCR

The identified 6 conserved miRNAs from transcriptomes were selected for validation and expression analysis through stem-loop qRT-PCR on the basis of their association with secondary metabolites biosynthesis in *P. hexandrum.* The transcript abundance of miRNAs ranged from 0.03 to 747-folds in roots and rhizomes of *P. hexandrum.* miR396b showed highest expression in roots with 2-fold and 747-fold in rhizomes of *P. hexandrum*. miR828b showed least expression in roots with 0.03-fold and miRf10132-akr 0.27-fold in rhizomes (Fig. 9).

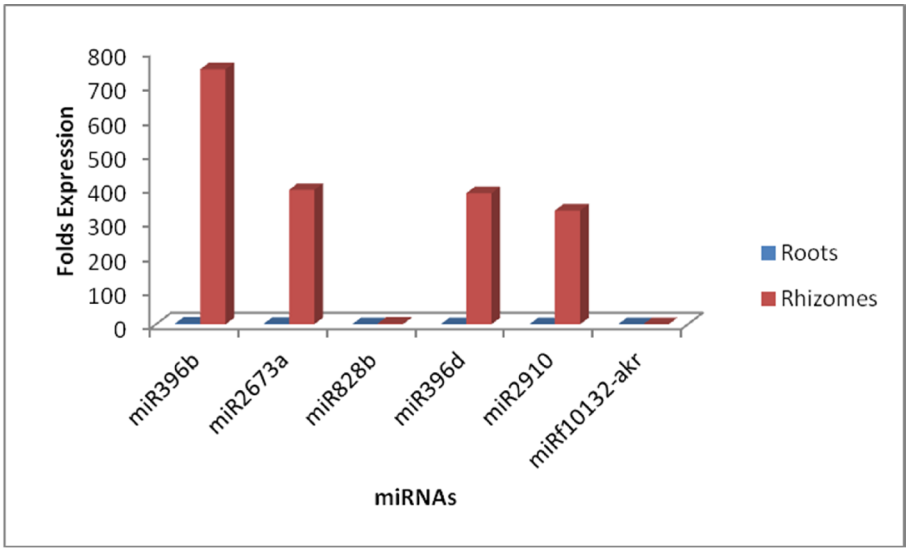

Fig. 9 Compartartive expression profile of miRNAs in roots and rhizomes w.r.t. shoots of *P. hexandrum*

Validation and expression analysis of miRNAs targets by qRT-PCR

Out of 375 predicted targets, eight targets (UDP glycosyltransferase (*UDP-Gly*), flavnol synthase (*FS*), glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*), peroxidise (*PRO*), malate dehydrogenase (*MD*), phosphoenolpyruvate carboxylase (*PEP*), WRKY and MYB transcription factors were selected for validation through qRT-PCR on the basis of their association with secondary metabolites biosynthesis. The expression profiles of target genes and their corresponding miRNAs could provide valuable information pertaining to regulation of secondary metabolites biosynthesis. The relative expression of target genes mediated by miRNAs ranged from 0.4 to 26 in roots and 0.1 to 28 in rhizomes of *P. hexandrum*. The target gene UDP-Gly showed higher transcript abundance with 305-fold in roots and WRKY gene with 130-fold in rhizomes of *P. hexandrum*. The PEP gene showed lowest expression with 0.4 fold and 0.1-fold in roots and rhizomes of *P. hexandrum* (Fig. 10).

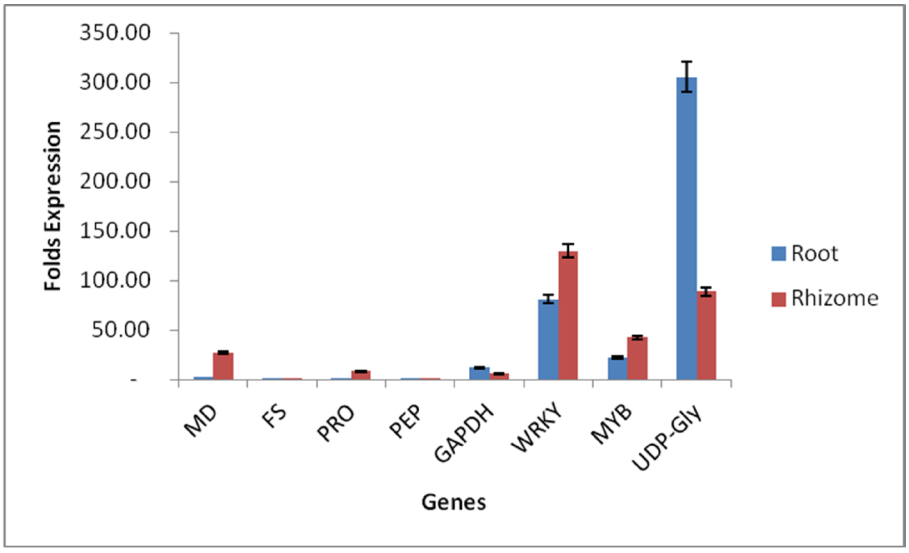

Fig. 10 Compartartive expression profile of miRNAs mediated target genes in roots and rhizomes w.r.t. shoots of *P. hexandrum*

Discussion

The demand for podophyllotoxin has been increasing day-by-day due to its usage in the manufacture of anti-cancer drugs. The limited raw material and low (0.3%) *in-vitro* production of podophyllotoxin in cell cultures necessitates the development of alternate production routes/sources to meet the industrial demands. Research on molecular components (pathway genes, TFs, ABC transporters and miRNAs) in plant species has opened the new window to enhance the metabolites content through understanding of the regulation mechanisms underlying in the biosynthesis of metabolites.

Validation and expression analysis of podophyllotoxin biosynthesis pathway genes

In order to gain better understanding on the molecular basis of podophyllotoxin biosynthesis, quantitative expression of 21 pathway genes was studied in different tissues of *P. hexandrum* by using two different approaches/methods (FPKM- based and qRT-PCR based expression) being employed nowadays. The FPKM values of pathway genes ranged from 1.23 to 118.43 in rhizomes transcriptomes of *P. hexandrum*. The qRT-PCR based relative transcript abundancy of pathway genes ranged from 0.004 to 23.06 in shoots, roots and rhizomes of *P. hexandrum* (Fig. 11). Transcript abundance/gene expression pattern observed through two

different platforms was found to be in conjunction with each other. These two methods/approaches was recently used for the validation of gene expression data in various plant species including *C. sinensis*, *S. tubersoum* and *Malus x domestica, etc.*, thereby, suggesting the accuracy of method [20, 21]. Out of 21 genes retrieved from GenBank Database (7 genes) and *P. hexandrum* transcriptomes (14 genes), 5 genes *SRD, C3H, CAD, CCR* and *C4H* showed higher expression in roots and rhizomes of *P. hexandrum* compared to shoots which was in agreement with the previous reports [22]. In addition, 4 genes *PD, CM*, *COMT* and *ADT* also showed higher expression (Fig. 11). The fold increase in transcript abundance was 4.98 and 8.04 for *PD*, 6.24 and 7.05 for *C3H*, 1.49 and 8.53 for *CM*, 4.35 and 7.90 for *COMT*, 2.37 and 4.48 for *CCR*, 3.05 and 8.98 for *C4H*, 5.38 and 13.76 for *CAD*, 6.88 and 8.44 for *ADT*, respectively in roots and rhizomes w.r.t. shoots of *P. hexandrum* (Fig. 11). The fold increase in transcript amount of genes were analysed by FPKM and qRT-PCR based approach in roots and rhizomes was consistent with the amount of podophyllotxoin content, which was 1.37% and 3.05% in roots and rhizomes, respectively. The expression of *SRD* gene was highest among all genes in the roots (19.46 folds) and rhizomes (23.05 folds) compared to shoots.

These obeservations indicated that podophyllotxoxin biosynthesis occurs in roots and rhizomes of *P. hexandrum* which is in accordance with the previous reports (Kumari et al., 2014). Our results were in concurrence with the previous reports where metabolite content was positively correlated with the expression pattern of biosynthetic pathway genes studied for picrosides in *P. kurroa*, catechins in *C. sinensis*, shikonin in *A. euchroma* and aconites in *A. heterophyllum* [23, 24].

The expression of *SRD* gene was highest among all genes in the roots (19.46 folds) and rhizomes (23.05 folds) compared to shoots. This gene is well known to catalyze secoisolariciresinol into matairesinol, a central precursor of podophyllotoxin biosynthesis [8]. Lan et al. [10] reported that *SRD* gene had maximum expression in root and rhizome tissues contributing to podophyllotoxin production in *Dysosma tsayuensis.* The transcript level of pathway genes such as *COMT, CM, PD, ADT, C4H, C3H, CCR, CAD* and *SRD* is positively correlated with the podophyllotoxin content as observed through Pearson's correlation coefficient values (≥0.99). Such values suggested a significant agreement between the transcript level and podophyllotoxin content. The higher expression of multiple genes in podophyllotoxin biosynthetic pathway may be due to the common regulation through regulatory components such as transcription factors, microRNAs (miRNAs) and *cis*-

regulatory elements. The expression of multiple genes of MIA pathway in *C. roseus* was attributed to a common regulator methyl jasmonate-responsive AP2 domain transcription factor, ORCA-3. Furthermore, the relatively higher expressions of *COMT, C4H, CCR, CAD* genes have also been reported in *Arabidopsis thaliana* for the production of lignin [25]. Previous studies have suggested that *C4H, CAD, CCR* play crucial role in the regulation of phenylpropanoid pathway leading to podophyllotoxin biosynthesis. Multiple genes of flavonoid pathway such as *C4H, CHS, CHI* and *FLS* showed increased transcript in relation to rutin content in *Fagopyrum* spp. Similarly, multiple genes such as *DXPS, ISPD, PMK* and *MECPS* showed elevated expression w.r.t. picrosides content in different tissues of *P. kurroa* [23]. Higher expression of multiple genes was also observed in rhizomes of ginger (*Zingiber officinale Rosc*.), turmeric (*Curcuma longa* L.), bamboo (*Phyllostachys praecox*), red rice (*Oryza longistaminata*) and *Sorghum propinquum* contributing to the biosynthesis of terpenoids [26].

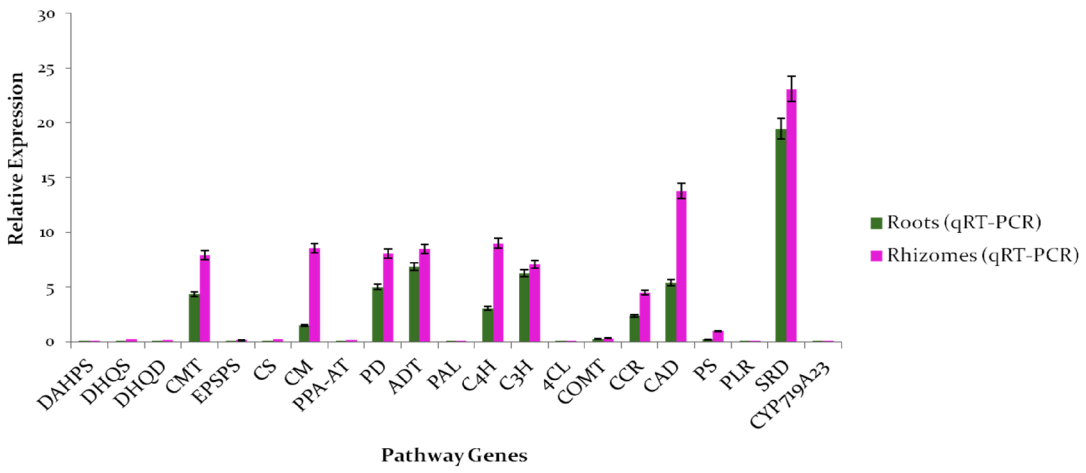

Fig. 11 Expression status of pathway genes in the roots and rhizomes w.r.t. shoots of *P. hexandrum*

Expression analysis of transcription factors vis-à-vis podophyllotoxin content

Podophyllotoxin biosynthesis occurs in different tissues of *Podophyllum* species through the contribution of phenylpropanoid/shikimic acid pathway which is controlled by regulatory components, particularly TFs. TFs are fascinating vital elements controlling the diverse physiological, biochemical and biological processes including growth, development, senescence, seed germination, environmental stress responses, etc. by modulating gene

expression. TFs are DNA-binding specific proteins that interact with *cis*-acting elements in the promoter regions of target genes, thus regulating the extent of gene expression [16]. It is crucial to know the status of TFs for better understanding the regulation of podophyllotoxin biosynthesis. Availability of transcriptomes of *Podophyllum* species provided a platform to identify and characterize candidate TFs families in shoots and rhizomes of *Podophyllum* species to infer their role in regulation of podophyllotoxin biosynthesis [9]. An extensive literature survey revealed that four TF families namely, bZIP, MYB, WRKY and bHLH regulate the phenylpropanoid/shikimic acid pathway in various plant species. These TF families might also be associated with the regulation of podophyllotoxin biosynthesis in *Podophyllum* species. Therefore, we focused on identifying these four TFs in shoots and rhizomes of *P. hexandrum* in reference to podophyllotoxin biosynthesis. We have mined transcriptomes of shoots and rhizomes of *P. hexandrum* in order to identify, scrutinize, validate and evaluate the transcript abundance of transcripts/genes that encode for bZIP, MYB, WRKY and bHLH TF families. Transcript abundance of respective TF families was evaluated by *in silico* FPKM and qRT-PCR-based transcript abundance methods. The FPKM values of TFs ranged between 0.0014-12.01 and 0.014-1162.01 in transcriptomes corresponding to shoots and rhizomes of *P. hexandrum* and *P. peltatum*, respectively. The qRT-PCR based relative transcript abundance of TF encoding transcripts ranged from 0.55 to 19.60-folds in rhizomes w.r.t. shoots of *P. hexandrum* (Fig. 7). Gene expression pattern observed through two different platforms was found to be in conjunction with each other (Fig. 6).

To ascertain whether the elevated levels of transcripts of TF families is only affecting the biosynthesis of podophyllotoxin in rhizomes/roots compared to shoots where podophyllotoxin content is almost negligible. We applied two-step refinement process where, the fold increase in transcript abundance of particular TFs was evaluated by comparing the transcripts abundance in rhizomes w.r.t shoots of *P. hexandrum* accessions varying for podophyllotoxin content. We have selected high (Ph5; 3.05%) and low (Ph6; 0.59%) podophyllotoxin content accession of *P. hexandrum* for comparative expression analysis. The transcripts corresponding to TFs families (MYB, WRKY, bHLH and bZIP) showed increase in transcript abundance with 1.09, 0.62, 0.26 and 19.60-folds in rhizomes compared to shoots of *P. hexandrum* (Fig. 7). To further confirm the expression status in relation to podophyllotoxin content, we compared the transcript abundance of TFs in rhizomes of high vs low podophyllotoxin content accessions of *P. hexandrum*. The expression of bZIP, MYB

and WRKY decreased to 0.04-15.78-folds between rhizomes of high versus low podophyllotoxin content. There was overall decrease in the expression level of transcription factors when compared between rhizomes of high versus low podophyllotoxin content accessions (Fig. 7). The transcript abundance of bZIP decreased from 19.60-fold to 3.82-fold. Similarly, the expression of MYB and WRKY decreased from 1.09-fold to 1-fold and 0.26-fold to 0.04-fold, respectively in rhizomes of high versus low podophyllotoxin content accessions. The differences in transcript abundance for identified transcription factors between high versus low podophyllotoxin content accessions reflects their association as per inherent genetic differences between accessions solely for podophyllotoxin contents. The transcript abundance of transcription factors was in conjunction with the metabolites content as revealed through comparative expression analysis between high versus low podophyllotoxin content accessions. Furthermore, a strong positive correlation was obtained between transcript abundance of TF families and podophyllotoxin content as observed by Pearson's correlation coefficient value (≥0.99). Expression profiling of genes in relation to metabolites content in high vs low metabolite content accessions of various plant species such as *Aconitum heterophyllum*, *Picrorhiza kurroa*, *Jatropha curcus*, etc. was studied previously, where metabolites contents were correlated with expression profiles [24, 27, 28].

The identified putative *cis*-regulatory elements in pathway genes are suggested to be involved in the regulation of podophyllotoxin biosynthesis pathway. The *in-silico* analysis revealed that binding sites for two transcription factors MYB and WRKY are commonly present in the promoter regions of pathway genes such as *CAD, C4H, PS, C3H, PAL, 4CL, PLR* and *SRD*. These transcription factors are suggested to regulate the podophyllotoxin biosynthesis by complementary binding with sequence elements in the promoters of target pathway genes. However, the validation of binding pattern of *cis*-regulatory elements with transcription factors can be taken up in future.

Experimental validation and expression profiling of ABC transporters

ABC transporters are known as largest families of transporters in plant systems having their role in transport of various macro and micro molecules across biological membranes. ABC transporters have broad substrate specificity and a number of ABC proteins have been implicated in transport of secondary metabolites in different plant species. Most of the identified transporters belonging to ABC family include ABCB/MDR (multi drug resistance), ABCC/MRP (multidrug and resistance protein) and ABCG/PDR (pleiotropic drug

resistance). The transporters have significant role in various physiological and biochemical processes in plant systems. The role of transporters is not fully understood due to limited availability of desired radiolabeled compounds requisite to decipher the transport mechanisms in plant systems. There is no report available on identification of ABC transporters genes and dissecting their role in secondary metabolites biosynthesis in *P. hexandrum*. Therefore, transcriptomes of *P. hexandrum* were mined to identify and validate ABC transporters genes.

252 ABC transporter genes were identified as unigenes in transcriptomes of *P. hexandrum*, which were further classified into 9 sub-families such as ABCA, ABCB, ABCC, ABCD, ABCE, ABCF, ABCG, ABCI and others in *P. hexandrum*. Phylogenetic analysis of ABC transporters showed high homology with the functionally characterised genes of *Nicotiana plumbaginifolia, Medicago truncatula, Arabidopsis thaliana*, *Zea mays,* etc. Clustering of ABC transporters genes with the well characterised genes of model plant species further confirmed the function of identified ABC transporters in *P. hexandrum.* Out of 252, 22 genes were selected for validation on the basis of transcript abundance to gain insights into regulation of podophyllotoxin biosynthesis. Transcript abundance for 22 ABC transporters genes was checked by two approaches- FPKM and qRT-PCR analysis. Expression analysis of 22 ABC transporters genes revealed 14 genes showing higher expression in shoots as compared to roots and rhizomes of *P. hexandrum* suggesting their major role in the transport of various macro and micromolecules in shoots. Shoots are the primary organs for synthesizing carbohydrates, starch and glucose in the photosynthesis required for various physio-biological and biochemical processes including secondary metabolites biosynthesis in plant systems [29]. The expression of ABC transporters in shoots might be due to the their involvement in transporting sugars and other carbohydrates, amino acids, peptides, polyamines, metal ions, etc., to rhizomes- the storage organ of *P. hexandrum*. The expression pattern observed by two approaches was found in concurrence with each other. Further in order to find out the suitable candidate genes correlating with podophyllotoxin content in *P. hexandrum,* Pearson's correlation coefficient was determined. Two genes, namely *PhABC6* and *PhABCIII* showed positive correlation with podophyllotoxin content. These genes were suggested to be the potential candidates for designing genetic intervention strategies aimed towards enhancement of podophyllotoxin content.

The podophyllotoxin biosynthesis occurs in roots/rhizomes of *P. hexandrum*; however the transport mechanisms of podophyllotoxin remain to be elucidated. Podophyllotoxin is mainly stored in the vacuole in rhizomes of *P. hexandrum*. ABC transporters were thought to be involved in transporting podophyllotoxin and its derivatives from different compartments to vacuoles of rhizomes. Several questions need to be answered such as how many transporters are involved in the transportation of podophyllotoxin. An attempt has been made by providing preliminary information of ABC transporters genes through identification and validation in different tissues/organs of *P. hexandrum.* ABC transporters were well known for transporting various secondary metabolites including lignins, flavonoids, terpenoids, etc. across plasma membranes and from cytosol to vacuole in the plant systems. To gain insights into podophyllotoxin biosynthesis and accumulation, we also evaluated the expression status of ABC transporters genes in roots and rhizomes with respect to shoots of *P. hexandrum.* Five genes namely *PhABC11*, *PhABC16*, *PhABC17*, *PhABCIII* and *PhABCIV* showed higher expression ranging from 4.45-folds to 20.53-folds in rhizomes and roots as compared to shoots. *PhABCIV* showed highest expression (20.53-folds) in rhizomes suggesting their major role in accumulation and transport of podophyllotoxin content in rhizomes of *P. hexandrum*. *PhABCIV* is well characterised for the transport of anthocyanidin 3-*O*-glucosides into the vacuole in *Vitis vinifera.*

The identified ABC transporters genes are thought to be involved in podophyllotoxin accumulation as their transcript abundance correlated with podophyllotoxin content. However, the functional validation of these highly correlated genes remains to be deciphered. The present study has set-up a new platform by providing ABC transporters gene sequences to scientific community for further research. This study would help in elucidating the physiological and biological functions of ABC transporters in *P. hexandrum*.

Role of miRNAs in regulation of podophyllotoxin biosynthesis

Our previous study revealed the presence of multiple pathway genes contributed towards podophyllotoxin biosynthesis [11]. In order to enhance the podophyllotoxin content in *in-vitro* the metabolic engineering requires at multiple steps of pathways or there is urgent need to find out the crucial regulatory components such as transcription factors and miRNAs involved in the regulation of podophyllotoxin. However, no information was available on molecular basis of regulation of podophyllotoxin biosynthesis in *P. hexandrum*.

The biosynthesis of primary and secondary metabolites controlled through the miRNA-mediated regulatory networks in plant systems. The NGS transcriptomes can provide valuable information about the role of miRNAs in various physio-biological processes in plants species. The present study identified for the first time 279 conserved miRNAs from transcriptomes known to regulate different biological processes in *P. hexandrum*. The putative mRNAs targets were identified for 279 conserved miRNAs using psRNATarget tool. This analysis revealed 896 mRNAs targets in *P. hexandrum* transcriptomes. To check the functions of identified 896 mRNA targets, KAAS analysis was performed which provides functional annotation for the mRNAs targets. 26 mRNAs targets have been implicated in different secondary metabolites biosynthetic pathways in plant species.

Out of 26 mRNAs targets, 8 targets *UDP-Gly, FS, GAPDH, PRO, MD, PEP, WRKY* and *MYB* transcription factors have been suggested to regulate the primary and secondary metabolites biosynthesis, including podophyllotoxin biosynthesis [30, 31, 22] (Table 2). Six miRNAs (miR396b, miR396d, miRf10132-akr, miR2673a, miR2910 and miR828b) were corresponded to these eight mRNAs targets (Table 2). miRNA, miR2673a/b had maximum number of targets (3 targets) i.e. peroxidise, MYB and WRKY transcription factors known to have their crucial role in controlling the biological processes as confirmed by psRNATarget and KAAS analysis.

Six miRNAs were further experimentally validated on the basis of their predicted mRNAs targets possibly involved in regulation of podophyllotoxin biosynthesis. qRT-PCR validation and expression profiling revealed candidate miRNAs from *P. hexandrum* that can act as suitable targets for designing a genetic intervention strategy. Out of several mRNA targets predicted, an important target is UDP glycosyltransferase governed by miR396b. The expression profiles of miR396b (747-fold) and its target *UDP-Gly* (~89-folds) was in negative correlation with each other in rhizomes of *P. hexandrum* showing downregulation with respect to podophyllotoxin content (3.02%). *UDP-Gly* played crucial role in the biosynthesis of secondary metabolites in various plant species through glucosylation of intermediates in biosynthetic pathways. *UDP-Gly* is known for its role in the formation of key intermediates of monoterpenoid and alkaloids biosynthesis pathways [32]). The involvement of UDP glycosyltransferase in podophyllotoxin biosynthesis has been shown previously, where it was suggested to be involved in the translocation of intermediates coumaryl and coniferyl alcohol in the form of glucosides from cytosol to cell wall [31].

In this study, we have also found one miRNA (miR2673a/b) targeting transcription factor (TF) MYB and WRKY (Table 2). MYB and WRKY TFs are known for their regulatory role in various physio-biological processes such as plant growth and development, responses to environmental cues, primary and secondary metabolites biosynthesis, etc. [33]. WRKY TFs (GaWRKY1, AaWRKY1, WRKY3, WRKY6 and WRKY33) have been implicated in the regulation of gossypol, artemisinin and camalexin biosynthesis in cotton, *Artemisia annua*, tobacco and *Arabidopsis*, respectively. Similarly, MYB TFs were involved in the regulation of physio-biological processes such as growth and development, metabolites biosynthesis, abiotic and biotic stress responses, seed formation and phenylpropanoid biosynthesis, etc. These TFs have also been suggested in the formation and regulation of podophyllotoxin biosynthesis [30]. The expression analysis showed negative correlation of MYB and WRKY TFs with their corresponding miRNAs in rhizomes of *P. hexandrum* with respect to podophyllotoxin content.

The selected miRNAs (miR396b, miR396d, miRf10132-akr, miR2673a, miR2910 and miR828b) showed possible role in the regulation of podophyllotoxin biosynthesis as their mRNA targets were implicated in different pathway modules such as glycolysis, TCA cycle, shikimic acid and phenylpropanoid pathways contributing towards podophyllotoxin biosynthesis (Fig. 7). The mRNA targets belonged to the biosynthesis of primary metabolites such as *GAPDH, PRO, MD* and *PEP* (Table 2). The biosynthesis of primary metabolites is prerequisite for the formation of secondary metabolites as the products of primary metabolic pathways act as precursors for the secondary metabolites biosynthesis, including podophyllotoxin biosynthesis (Fig. 7). The expression analysis of *GAPDH, PRO, MD* and *PEP* and their miRNAs revealed negative correlation thereby, suggesting the downregulation of these mRNA targets by complementary miRNA:mRNA base pairing in *P. hexandrum* with respect to podophyllotoxin content (Fig. 6). Due to downregulation of these mRNA targets the biosynthesis of podophyllotoxin got affected and ultimately lead to lesser podophyllotoxin production. The identified miRNAs could be knock-down to increase the podophyllotoxin content in *P. hexandrum.* The selected six miRNAs have been known in many plant species for their role in regulation of metabolites such as miR396b was regulating the growth regulating factor (GRF) and (basic helix-loop-helix) bHLH family of transcription factors in *Arabidopsis thaliana* [34]. Similarly, miR828b was shown to regulate the MYB transcription factor involved in the biosynthesis of flavonoid and lignin biosynthesis in *Arabidopsis thaliana* [35].

The identified miRNAs in the present study could be suitable targets for designing a genetic intervention strategy (knock-down/antisense) to enhance the podophyllotoxin content in *P. hexandrum*. The present study is the first report on identification, characterisation and detailed expression profiling of miRNAs and their targets in *P. hexandrum.*

Conclusion and future prospects

The outcome of this work results in the identification and validation of crucial molecular components (pathway genes, TFs, ABC transporters and miRNAs) contributing to podophyllotoxin machinery in *P. hexandrum.* The information on molecular basis of these molecular components will provide further insights into the existing knowledge of podophyllotoxin biosynthesis. Expression analysis revealed nine genes (*SRD, PD, C3H, CM, COMT, CAD, CCR, C4H* and *ADT*) of podophyllotoxin biosynthetic pathway in root and rhizome tissues of *P. hexandrum* that were positively correlated with the podophyllotoxin content. Further, the work on regulatory components i.e. TFs and miRNAs revealed four key transcription factors (MYB, WRKY, bHLH and bZIP) and six miRNAs (miR396b, miR396d, miRf10132-akr, miR2673a, miR2910 and miR828b) targeted UDP glycosyltransferase, flavnol synthase, glyceraldehyde 3-phosphate dehydrogenase, peroxidise, malate dehydrogenase, phosphoenolpyruvate carboxylase, WRKY and MYB transcription factors showed possible role in regulation of podophyllotoxin biosynthesis in *P. hexandrum.* These regulatory components would facilitate the research on understanding the regulatory machinery of podophyllotoxin biosynthesis. Moreover, the work on ABC transporters would lay a basis for understanding the transportation of podophyllotoxin by revealing the candidate transporters genes *PhABC6* and *PhABCIII* that showed positive correlation with podophyllotoxin content. Overall, the present work deciphered molecular components involved in the biosynthetic machinery of podophyllotoxin that could be the key targets for designing a genetic intervention strategy to enhance podophyllotoxin content followed by functional validation using approaches like enzyme assays and gene silencing. Moreover, the current work opens a new window for future endeavours to further exploration and research in *P. hexandrum* at a following link: https://sites.google.com/site/combiogroup/data-access.

References

[1] D.E Jackson. P.M. Dewick, "Aryltetralin lignans from *Podophyllum hexandrum* and *Podophyllum peltatum*", *Phytochemistry*, vol. 23, pp.1147-1152, 1984.

[2] A. Nag, P. Bhardwaj, P.S. Ahuja, and R.K. Sharma, "Identification and characterization of novel UniGene-derived microsatellite markers in *Podophyllum hexandrum* (Berberidaceae)", *J. Genet,* vol. 92, pp. e4–e7, 2013

[3] O. Pagani, M. Zucchetti, C. Sessa, J. Jong, M.D. Incalci, M.D. Fusco, A. Kaeser-Fröhlich, A. Hanauske, and F. Cavalli, "Clinical and pharmacokinetic study of oral NK 611, a new podophyllotoxin derivative", *Cancer Chemotherapy Pharmacol*, vol. 38, pp. 541–547, 1996.

[4] R. Gupta, and K.L. Sethi, "Conservation of medicinal plant resources in the Himalayan region. In: Conservation of tropical plant resources", Eds. SK Jain and KL Mehra. Botanical survey of India. *Howarh,* pp. 101-107, 1983.

[5] D. Bhattacharyya , R. Sinha, S. Ghanta , A. Chakraborty , S. Hazra, and S. Chattopadhyay, "Proteins differentially expressed in elicited cell suspension culture of *Podophyllum hexandrum wi*th enhanced podophyllotoxin content", *Proteome Sci,* vol. 10, pp.34. doi: 10.1186/1477-5956-10-34, 2012.

[6] A. Giri, L.M. Narasu, "Production of podophyllotoxin from *Podophyllum hexandrum*: a potential natural product for clinically useful anticancer drugs", *Cytotechnology*, vol.34, pp. 17–26, 2000.

[7] S. Chattopadhyay, R.S. Mehra, A.K. Srivastava, S.S. Bhojwani, and V.S. Bisaria, "Effect of major nutrients on podophyllotoxin production in *Podophyllum hexandrum* suspension cultures", *Applied Microbiology and Biotechnology,* vol. 60, pp. 541-546, 2003.

[8] Z.Q. Xia, M.A. Costa, J. Proctor, L.B. Davin, and N.G. Lewis, "Dirigent-mediated podophyllotoxin biosynthesis in *Linum flavum* and *Podophyllum peltatum*", *Phytochemistry* vol. 55, pp. 537‑549, 2000.

[9] J.V. Marques, K.W. Kim, C. Lee, M.A. Costa, G.D. May, J.A. Crow, L.B. Davin, and N.G. Lewis, "Next generation sequencing in predicting gene function in podophyllotoxin biosynthesis", *Journal of Biological Chemistry*, vol. 288, pp. 466-479, 2013.

[10] X. Lan, S. Ren, Y. Yang, M. Chen, C. Yang, H. Quan, G. Zhong, and Z. Liao, "Molecular cloning and characterization of the secoisolariciresinol dehydrogenase gene

involved in podophyllotoxin biosynthetic pathway from Tibet *Dysosma"*, *Journal of Medicinal Plants Research,* vol. 4, pp. 484-489, 2010.

[11] P. Kumar, T. Pal, N. Sharma, V. Kumar, H. Sood, and R.S. Chauhan, "Expression analysis of biosynthetic pathway genes vis-à-vis podophyllotoxin content in *Podophyllum hexandrum* Royle", *Protoplasma,* vol. 252(5), pp. 1252-1262, 2015.

[12] K. Tamura, G. Stecher, D. Peterson, A. Filipski, and S. Kumar, "MEGA6: Molecular Evolutionary Genetics Analysis version 6.0", *Molecular Biology and Evolution*, vol. 30, pp. 2725-2729, 2013.

[13] X. Dai, and P.X. Zhao, "PsRNATarget: a plant small RNA target analysis server", *Nucleic Acids Res*, vol. 39, pp. 155 –159, 2011.

[14] B. Li and C.N. Dewey, "RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome", *BMC Bioinformatics,* vol. 12, pp. 323, 2011.

[15] K. Higo, Y. Ugawa, M. Iwamoto, and T. Korenaga, "Plant *cis*-acting regulatory DNA elements (PLACE) database", *Nucleic Acids Res*, vol. 27, pp. 297-300, 1999.

[16] C.Q. Yang, X. Fang, X.M. Wu, Y.B. Mao, L.J. Wang, and X.Y. Chen, "Transcriptional regulation of plant secondary metabolism", *J Integr Plant Biol*, vol. 54, pp. 703-712, 2012.

[17] Xu D.Q., Huang J., Guo S.Q., Yang X., Bao Y.M., Tang H.J., and H.S. Zhang, "Over expression of a TFIIIA-type zinc finger protein gene ZFP252 enhances drought and salt tolerance in rice (*Oryza sativa* L.)", *FEBS Letters*, vol. 582, pp. 1037–1043, 2008.

[18] Z. Zhang, J. Yu, D. Li, Z. Zhang, F. Liu, X. Zhou, T. Wang, Y. Ling and Z. Su, "PMRD: plant microRNA database", Nucleic acids research, vol. 3, pp. D806-D813, 2010.

[19] Y. Moriya, M. Itoh, S. Okuda, A. Yoshizawa, and M. Kanehisa, "KAAS: an automatic genome annotation and pathway reconstruction server", Nucleic Acids Res, vol. 35, W182-W185, 2007.

[20] A. Paul, S. Jha, S. Bhardwaj, S. Singh, Shankar R., and S. Kumar, "RNA-seq-mediated transcriptome analysis of actively growing and winter dormant shoots identifies non-deciduous habit of evergreen tree tea during winters", *Scientific Reports*, vol. 4, pp. 5932, doi: 10.1038/srep05932, 2014.

[21] A.K. Singh, V. Sharma, A.K. Pal, V. Acharya, and P.S. Ahuja, "Genome-wide organization and expression profiling of NAC transcription factor family in potato (*Solanum tuberosum* L)", *DNA research*, Vol. 20, pp. 403-423, 2013.

[22] D. Bhattacharyya, R. Sinha, R. Datta, S. Hazra, and S. Chattopadhyay, "De novo transcriptome analysis using 454 pyrosequencing of the Himalayan May apple, *Podophyllum hexandrum", BMC Genomics*, vol.1, pp. 748, 2013.

[23] S. Pandit, K. Shitiz, P.K. Naik, H. Sood, and R.S. Chauhan, "Expression pattern of fifteen genes of non-mevalonate (MEP) and mevalonate (MVA) pathways in different tissues of endangered medicinal herb *Picrorhiza kurroa* with respect to Picrosides Content", *Molecular Biology Reports*, vol. 40, pp. 1053-63, 2013.

[24] N. Malhotra, V. Kumar, H. Sood, T.R Singh, and R.S. Chauhan, "Multiple genes of mevalonate and non-mevalonate pathways contribute to high aconites content in an endangered medicinal herb, *Aconitum heterophyllum* Wall, *Phytochemistry*, vol. 108, pp. 26-34, 2014.

[25] M.B Ali, and D.H. McNear, "Induced transcriptional profiling of phenylpropanoid pathway genes increased flavonoid and lignin content in *Arabidopsis* leaves in response to microbial products", *BMC Plant Biol*, vol. 1, pp.84. doi: 10.1186/1471-2229-14-84, 2014.

[26] K.H. Wang, H.Z. Peng, E.P. Lin, Q.Y. Jin, X.Q. Hua, S. Yao, H.W. Bian, N. Han, and J.W. Pan, "Identification of genes related to the development of bamboo rhizome bud", *J Exp Bot,* vol. *6*1, pp. 551–561, 2010.

[27] K. Shitiz, N. Sharma, .T Pal, H. Sood, and R.S. Chauhan, "NGS transcriptomes and enzyme inhibitors unravel complexity of picrosides biosynthesis in *Picrorhiza kurroa* Royle ex. Benth", *Plos One*, 10(12): e0144546, 2015.

[28] Sood, R.S. Chauhan, "Regulation of FA and TAG biosynthesis pathway genes in endosperms and embryos of high and low oil content genotypes of *Jatropha curcas* L", *Plant Physiology and Biochemistry,* vol. 94, pp. 253-267, 2015

[29] S.C. Huber, "Biochemical mechanism for regulation of sucrose accumulation in leaves during photosynthesis", *Plant Physiology*, vol. 91, pp. 656-662, 1989.

[30] P. Kumar, V. Jaiswal, T. Pal, J. Singh, and R.S. Chauhan, "Comparative whole-transcriptome analysis in *Podophyllum* species identifies key transcription factors contributing to biosynthesis of podophyllotoxin in *P. hexandrum*", *Protoplasma* 10.1007/s00709-015-0938-7, 2016.

[31] A. Kumari, H.R. Singh, A. Jha, M.K. Swarnkar, R. Shankar, and S. Kumar, "Transcriptome sequencing of rhizome tissue of *Sinopodophyllum hexandrum* at two temperatures", *BMC Genomics*, vol. 7 pp. 871, 2014.

[32] A. Pani, and R.K. Mahapatra, “Computational identification of microRNAs and their targets in Catharanthus roseus expressed sequence tags”, *Genomics Data*, vol. 1, pp. 2-6, 2013T.

[33] C. Dubos, R. Stracke, E. Grotewold, B. Weisshaar, C. Martin, L. Lepiniec, “MYB transcription factors in Arabidopsis”, *Trends Plant Sci*, vol. 15, pp. 573–581, 2010.

[34] J.M. Debernardi, R.E. Rodriguez, M.A. Mecchia, and J.F. Palatnik, “Functional specialization of the plant miR396 regulatory network through distinct microRNA-target interactions”, *PLoS Genet*, vol. 8, pp. e1002419, doi:10.1371/journal.pgen.1002419, 2012.

[35] R. Rajagopalan, H. Vaucheret, J. Trejo, and D.P. Bartel, “A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*”, *Genes Dev*. Vol. 20, pp. 3407-25, 2006.

List of Publications

Papers in Peer Reviewed Journals

1. **Pawan Kumar,** Tarun Pal, Neha Sharma, Varun Kumar, Hemant Sood, Rajinder S. Chauhan (2015) Expression analysis of biosynthetic pathway genes vis-à-vis podophyllotoxin content in *Podophyllum hexandrum* Royle. Protoplasma 252(5): 1253-1262 **(IF- 2.65)**

2. **Pawan Kumar**, Varun Jaiswal, Tarun Pal, Jagdish Singh, Rajinder S. Chauhan (2016) Comparative whole-transcriptome analysis in *Podophyllum* species identifies key transcription factors contributing to biosynthesis of podophyllotoxin in *P. hexandrum*. Protoplasma 10.1007/s00709-015-0938-7 **(IF-2.65)**

3. **Pawan Kumar**, Ritika Sharma, Varun Jaiswal, Rajinder S. Chauhan (2016) Transcriptome-wide identification, validation and expression analysis of ABC transporters for dissecting their role in podophyllotoxin biosynthesis in *Podophyllum hexandrum* Royle. Biologia Planterum **(IF-1.85)** (In Press)

4. **Pawan Kumar**, Jibesh Kumar Padhan, Ashwani Kumar, Rajinder S. Chauhan (2016) Identification of miRNAs in transcriptomes of *Podophyllum hexandrum* and their association with the biosynthesis of secondary metabolites. Plant Cell Reports **(IF-3.02)** (Under Review)

5. Tarun Pal*****, Jibesh Kumar Padhan*****, **Pawan Kumar**, Hemant Sood, Rajinder S. Chauhan (2016) Comparative transcriptomics uncovers differences in physiological/biochemical processes between photoautotrophic and photoheterotrophic modes of nutrition vis-à-vis secondary metabolites biosynthesis in *Swertia chirayita*. Plant Cell, Tissue and Organ Culture **(IF-2.71)** (Under Review)

6. Jibesh Kumar Padhan, **Pawan Kumar**, Hemant Sood, Rajinder S. Chauhan (2016) Prospecting NGS-Transcriptomes to assess regulation of miRNA-inspected secondary metabolites biosynthesis in *Swertia chirayita*, a medicinal herb of North-Western Himalayas. Molecular Biology Reports **(IF-2.02)** (Submitted)

Conference Papers/Abstracts

1. **Pawan Kumar**, Tarun Pal, Varun Jaiswal, Hemant Sood, Rajinder S. Chauhan (2014). Computational mining of transcriptomes for transcription factors controlling various

biological processes in *Podophyllum* spp. Agrotechnology; 2:4, http://dx.doi.org/10.4172/2168-9881.S1.008

2. **Pawan Kumar**, Neha Sharma, Tarun Pal, Hemant Sood, Rajinder S. Chauhan (2014). Differential expression analysis of phenylpropanoid pathway genes controlling podophyllotoxin biosynthesis in *Podophyllum hexandrum*: A medicinal herb of North-Western Himalayas. Proceedings of the National Symposium on Advances in Biotechnology for Crop Improvement , Baru Sahib, Himachal Pradesh, India July 12, 2014

3. **Pawan Kumar**, Neha Sharma, Tarun Pal, Hemant Sood, Rajinder S. Chauhan (2014). Molecular components control podophyllotoxin biosynthesis in *Podophyllum hexandrum*: A medicinal herb of North-Western Himalayas. Proceedings of International Conference on Emerging Trends in Biotechnology, page no-14, Jawaharlal Nehru University, New Delhi, India, 6th Nov-9th Nov. 2014

4. Neha Sharma, **Pawan Kumar**, Tarun Pal, Hemant Sood, Rajinder S. Chauhan (2014). In-vitro conservation and SSR fingerprinting in *Podophyllum hexandrum* Royle accession from North-Western Himalayas National Symposium on Advances in Biotechnology for Crop Improvement, Eternal University, Baru Sahib, Himachal Pradesh, India July 12, 2014

Best Poster Award

- **Pawan Kumar**, Tarun Pal, Varun Jaiswal, Hemant Sood and Rajinder S. Chauhan (2014) Computational mining of transcriptomes for transcription factors controlling various biological processes in *Podophyllum* spp. Poster presented in 2nd International Conference on Agricultural and Horticultural Sciences (Omics Group) Radisson Blu Plaza Hotel, Hyderabad, India, February 03-05, 2014

NCBI GenBank Submissions

Kumar P, Sharma N, Sood A, Sood H, and Chauhan RS (2015):

- Partial cloned sequence of MYB Family transcription factor (*PhMYB*) mRNA, partial cds KP299959
- Partial cloned sequence of bZIP Family transcription factor (*PhbZIP*) mRNA, partial cds KP299960
- Partial cloned sequence of WRKY Family transcription factor (*PhWRKY*) mRNA, partial cds KP299961